**Journal of Water and Soil**
**Homepage:** https://jsw.um.ac.ir

**Research Article**
**Vol. 37, No. 6, Feb.-Mar., 2024, p. 923-942**

# Digital Mapping of Soil Texture Particles with Machine Learning Models and Environmental Covariates

**P. Khosravani [1], M. Baghernejad [2*], A.A. Moosavi[2], S.R. FallahShamsi[3]**

1 and 2- Department of Soil Science, College of Agriculture, Shiraz University, Shiraz, Iran
(*- Corresponding Author Email: majidbaghernejad@yahoo.co.uk)
3- Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran

| | |
|---|---|
| Received: 12-09-2023<br>Revised: 13-11-2023<br>Accepted: 26-11-2023<br>Available Online: 27-11-2023 | **How to cite this article:**<br>Khosravani, P., Baghernejad, M., Moosavi, A.A., & FallahShamsi, S.R. (2024). Digital mapping of soil texture particles with machine learning models and environmental covariates. *Journal of Water and Soil, 37*(6), 923-942. (In Persian with English abstract). https://doi.org/10.22067/jsw.2023.84413.1331 |

## Introduction

Understanding the particle size distribution (PSD) is of great importance for plant growth and soil management. In recent years, the science of soil has witnessed a significant increase in digital soil mapping (DSM) activities. In this regard, machine learning models (ML) have emerged as an alternative and tool for DSM, which are mainly used for data mining and pattern recognition purposes, and are now widely used for regression and classification tasks in all fields of science. Hence, this study was undertaken to spatially model sand, silt, and clay particles utilizing machine learning models such as Random Forest (RF), Support Vector Regression (SVR), and the Co-Kriging geostatistical model. Additionally, auxiliary variables with high spatial resolution were incorporated into the analysis. This investigation was conducted in a section of the Marvdasht plain, located in Fars province.

## Materials and Methods

The present study was conducted in a part of Marvdasht plain located between 35.82´41°52' to 1.07´57°52' east longitude and 35.02´48°29' to 14.72´2°30' north latitude, and 40 km north of Shiraz with an area of about 50,000 hectares. After determining the study area boundaries, the positions of 200 sampling points were determined using the R software and the conditioned Latin hypercube sampling method. In other words, for soil feature modeling, 200 samples were taken from two depths of zero to 30 and 30 to 60 centimeters in the study area. Then, the samples were transferred to the laboratory, dried, and passed through a 2 mm sieve. Finally, the soil texture components were measured by the hydrometer method. The environmental variables used in this study are a wide range of representatives of soil-forming factors that were prepared as much as possible from sources with minimum cost and high accessibility. In total, 75 environmental variables were prepared, and the raster format related to all environmental variables, including 39 elevation and altitude variables and 36 remote sensing measurement variables, was extracted. Finally, the factor-tuning inflation variance and Boruta algorithm were used to select the optimal variables.

## Results

The minimum amount of clay was measured at 10.21% and 10.45%, respectively, and the maximum amount was 32.65% and 36.35% at the surface and subsurface depths. The average amount of clay in all samples was 37.91% and 35.61%. The average amount of sand was measured at 25.65% and 26.02% at the surface and subsurface depths, respectively. The maximum amount of sand was observed in the northern and higher parts of the study area and was



https://doi.org/10.22067/jsw.2023.84413.1331

equal to 54.68% and the minimum amount was predicted in the low-lying areas of the study area. Low-lying areas and sedimentary plains in the central part of the study area contained high amounts of silt. Four depth variables valley depths (VD), texture (TE), topographic wetness index (TWI), and clay index (CI) related to geomorphometric parameters and the normalized difference vegetation index (NDVI) variable related to remote sensing indices were selected as optimal variables. The RF model with $R^2$ of 54.0% and 36.0% for predicting sand, 48.0% and 64.0% for predicting silt, and 52.0% and 49.0% for predicting clay at both surface and subsurface depths performed better than the SVR and Co-Kriging models. The most effective variable in predicting the spatial distribution of soil particles was VD with relative importance of 60% and 65% for predicting sand at the surface and subsurface depths, 70% for predicting silt at the surface depth, and 70% and 65% for predicting clay at both surface and subsurface depths, respectively. Only TE and TWI variables were more important than VD for predicting silt at subsurface depth. These results show that topographic variables are effective in the spatial variation of soil particles. Unlike clay, the highest amount of sand in both depths was observed in the northern part and the highest part of the study area, and the lowest amount was predicted in the low-lying areas of the study area.

## Conclusion

In general, with the aim of this research, maps of the spatial distribution of soil texture components were prepared at both surface and subsurface depths using machine learning and geostatistical approaches along with environmental covariates in a part of Marvdasht plain. Among the selected environmental covariates, topographic attributes, especially the valley depth (VD), had the highest effect in justifying the spatial prediction of soil texture components. Also, the results of comparing the performance of machine learning models supported the higher efficiency of the RF model than other models. Therefore, the approach used in this study to prepare a map of soil texture components can be useful as a guide for mapping useful soil features in areas with similar climatic and topographic conditions.

**Keywords:** Co-Kriging, Digital soil mapping, Machine learning, Random Forest, Soil particle size distribution

**نشریه آب و خاک**
https://jsw.um.ac.ir

**مقاله پژوهشی**

جلد ۳۷، شماره ۶، بهمن-اسفند ۱۴۰۲، ص. ۹۴۲-۹۲۳

# نقشه‌برداری رقومی اجزای بافت خاک با استفاده از رویکردهای یادگیری ماشین و متغیرهای محیطی

**پگاه خسروانی[۱]- مجید باقرنژاد[۲*]- سید علی اکبر موسوی[۲]- سید رشید فلاح شمسی[۳]**

تاریخ دریافت: ۱۴۰۲/۰۶/۲۱
تاریخ پذیرش: ۱۴۰۲/۰۹/۰۵

## چکیده

بافت به‌عنوان یکی از ویژگی‌های موثر بر وضعیت هیدرولوژیکی و مکانیکی خاک از اهمیت ویژه‌ای برای رشد گیاهان برخوردار است. بنابراین، مطالعه حاضر با هدف مدل‌سازی مکانی ذرات خاک شامل شن، سیلت و رس با استفاده از مدل‌های یادگیری ماشین جنگل تصادفی (RF)، رگرسیون بردار پشتیبان (SVR) و مدل زمین‌اماری کوکریگینگ (Co-Kriging) و نقشه مدل رقومی ارتفاع و تصاویر ماهواره‌ای با وضوح مکانی بالا در دشت مرودشت استان فارس انجام گرفت. مطالعات میدانی شامل حفر ۲۰۰ پروفیل مشاهداتی براساس روش نمونه‌برداری ابر مکعب لاتین (cLHS) و نمونه‌برداری از دو عمق، سطحی (۳۰ سانتی‌متر) و زیرسطحی (۶۰-۳۰ سانتی‌متر) بود. طیف وسیعی از متغیرهای کمکی با دسترسی آسان شامل پارامترهای توپوگرافی و شاخص‌های سنجش از دور به‌عنوان پیش‌بینی کننده‌های محیطی با وضوح مکانی ۳۰ متر تهیه گردید. مناسب‌ترین متغیرهای کمکی با استفاده از آزمون هم‌خطی توسط شاخص تورم واریانس (VIF) و سپس رویکرد انتخاب متغیر باروتا انتخاب شد. نتایج نشان داد که مدل RF نسبت به مدل‌های SVR و Co-Kriging دارای عملکرد بهتر و دقت بیشتر برای پیش‌بینی هر سه ذره شن، سیلت و رس است. ضریب تبیین ($R^2$) پیش‌بینی شده برای رس در عمق سطحی و زیر سطحی به‌ترتیب ۰/۵۲ و ۰/۴۹، برای سیلت ۰/۴۸ و ۰/۶۴ و برای شن ۰/۵۴ و ۰/۳۶ درصد محاسبه شد. همچنین متغیر عمق دره مهمترین متغیر در پیش‌بینی توزیع مکانی شن، سیلت و رس شناسائی شد. تهیه نقشه توزیع مکانی اندازه ذرات خاک حاصل می‌تواند منبع مهمی از اطلاعات برای مدیریت کارآمد منابع زمین و نظارت بر محیط‌زیست باشد. به‌طور کلی رویکرد مورد استفاده در این تحقیق به‌دلیل استفاده از منابع اطلاعاتی با دسترسی آسان و ارزان می‌تواند به‌عنوان راهنمایی برای تهیه‌ی نقشه‌های مدیریت‌پذیر خاک در دشت‌های کشاورزی کشور با قابلیت به‌روز رسانی در حداقل زمان ممکن مورد استفاده قرار گیرد.

**واژه‌های کلیدی:** توزیع اندازه ذرات خاک، جنگل تصادفی، کوکریجینگ، نقشه‌برداری رقومی، یادگیری ماشین

## مقدمه

بافت خاک به‌عنوان یکی از شاخص‌های مهم کیفیت خاک بر روی بسیاری ازخصوصیات فیزیکی، شیمیایی، بیولوژیکی، هیدرولوژیکی و مکانیکی خاک مؤثر است (Hossain *et al.*, 2018). آگاهی از توزیع اندازه ذرات خاک (PSD)[۳] برای رشد گیاه و مدیریت خاک از اهمیت بالایی برخوردار است (Parent *et al.*, 2021). بنابراین، به‌طور مستقیم کیفیت خاک را کنترل کرده و برای تصمیم‌گیری‌های مدیریتی بسیار پرکاربرد می‌باشد (de Jesus Duarte *et al.*, 2019). PSD همچنین نقش کلیدی برای درک پیدایش و توزیع خاک ایفا کرده و با استفاده از آن می‌توان درجه پدوژنتیک و نرخ هوازدگی را تخمین زد (Lucas *et al.*, 2019). تغییرپذیری عمودی و افقی بافت خاک توسط بسیاری از

۱ و ۲- گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شیراز، شیراز، ایران
(*- نویسنده مسئول: Email: majidbaghernejad@yahoo.co.uk)
۳- گروه مهندسی منابع طبیعی و محیطزیست، دانشگاه شیراز، شیراز، ایران

https://doi.org/10.22067/jsw.2023.84413.1331

3- Particle Size Distribution

فرآیندها و عوامل مرتبط با پدوژنز مانند جابه‌جایی عمودی ذرات [۱] در پروفیل خاک، حذف سطحی، تجمع و مواد مادری کنترل می‌شود (Zhang *et al.*, 2023).

علی‌رغم اهمیت ویژگی‌های خاک مرتبط با PSD، کمبود اطلاعات مکانی درمورد آن‌ها در بسیاری از کاربری‌های اراضی وجود دارد (Chen *et al.*, 2021). بنابراین، به‌روز رسانی و گسترش پایگاه‌های داده PSD خاک که به‌طور خاص برای برنامه‌های مدل‌سازی طراحی شده‌اند، حیاتی است. در سال‌های اخیر، علم خاک شاهد افزایش قابل توجهی در فعالیت‌های نقشه‌برداری رقومی خاک (DSM)[۲] بوده است. این امر ناشی از همگرایی چندین عامل همزمان از جمله، تقاضای زیادی برای اطلاعات کمی و مکانی خاک، انباشتگی پایگاه‌های اطلاعاتی ویژگی‌های خاک اندازه‌گیری شده همراه با متغیرهای محیطی کاملاً شناخته‌شده، و توسعه مدل‌های عددی ترکیبی با استفاده از منابع کامپیوتری برای استخراج این ذخایر داده‌های خاک است (Arrouays *et al.*, 2020).

چارچوب DSM با توسعه‌ی مدل فاکتوریل ینی (Jenny, 1941) توسط مک براتنی و همکاران (McBratney *et al.*, 2003) با عنوان مدل اسکورپن[۳] رسمیت یافت، که در آن با استفاده از روش‌های آماری و ریاضیاتی مختلف و کمی‌سازی متغیرهای محیطی (نمایندگان فاکتورهای خاکسازی) اقدام به پیش‌بینی مکانی خصوصیات یا کلاس-های خاک می‌شود.

به‌طور متعارف، پیش‌بینی مکانی خاک در چارچوب مدل‌های زمین‌آماری تعبیه شده است (Ma *et al.*, 2019) که در آن نمونه‌ای از یک ویژگی خاک به‌عنوان مجموع ترکیب خطی متغیرهای محیطی و یک باقیمانده همبسته (تصادفی) مدل‌سازی می‌شود. مدل‌های زمین آماری اغلب در نقشه‌برداری خاک مورد استفاده قرار می‌گیرند زیرا دارای مزایای متعددی هستند (Radočaj *et al.*, 2022). اول، امکان تفسیر فرآیندهای فیزیکی زیربنایی ارائه شده توسط مدل را فراهم می‌کند. ثانیا، مدل‌سازی با توجه به همبستگی مکانی داده‌ها انجام می‌شود. این ویژگی برای متغیرهایی مانند خاک که از مکانی به مکان دیگر متفاوت است اما بین مکان‌ها همبستگی نشان می‌دهد، بسیار حائز اهمیت می باشد. سوم امکان اندازه‌گیری عدم قطعیت مدل‌سازی وجود دارد. در بسیاری از شرایط مانند یک تصمیم‌گیری پیش‌بینی تنها معیار نیست و نیاز به نقشه‌های عدم قطعیت می‌باشد. نقشه‌برداری زمین آماری خاک، همچنین، دارای محدودیت‌های متعددی است. به‌عنوان مثال داده‌ها باید دارای توزیع نرمال، ثابت (میانگین و واریانس با واحد ثابت) و همسانگرد[۴] باشند (Paramasivam, 2019). همچنین مدل‌سازی رابطه غیرخطی بین یک ویژگی یا کلاس خاک و متغیرهای کمکی متقابل متعدد ساده نیست و چالش‌های بیشتری دارد. در مناطق ناهمگن، مدل تغییرات مکانی (یعنی واریوگرام) نیز نمی‌تواند تغییرات تدریجی و ناگهانی در تغییرات خاک را ثبت کند. در نهایت، اگر حجم نمونه و یا تعداد مکان‌های پیش‌بینی زیاد باشد، مدل‌های زمین‌آماری از لحاظ محاسباتی ضعیف هستند (Paramasivam, 2019). بنابراین در اغلب موارد داده‌های خاک و تقاضای کاربر، انطباق با این مفروضات را به یک کار چالش برانگیز تبدیل می‌کند.

رویکردهای یادگیری ماشین (ML)[۵]، به‌عنوان ابزاری مفید برای DSM می‌باشند(Padarian *et al.*, 2019)، و تکنیک‌های ML به دسته بزرگی از الگوریتم‌های مبتنی بر داده‌های غیرخطی اشاره می‌کنند که عمدتاً برای اهداف داده‌کاوی و تشخیص الگو به کار می‌روند، و اکنون اغلب برای کارهای رگرسیون و طبقه‌بندی در همه زمینه‌های علم استفاده می‌شوند (Wang *et al.*, 2020). بر خلاف روش‌های زمین آماری که معمولاً برای برآورده کردن مفروضات تغییر مشاهدات اصلی لازم است، الگوریتم‌های ML فرضی را از توزیع مشاهدات ایجاد نمی‌کنند. الگوریتم‌های ML همچنین می‌توانند تعداد زیادی متغیرهای همبسته متقابل را به‌عنوان پیش‌نمایش اداره کنند (Wadoux *et al.*, 2020). با ظهور الگوریتم‌های ML مبتنی بر داده در علم خاک پیش‌بینی ویژگی‌های خاک مانند مدل‌سازی طبقات خاک (Poppiel *et al.*, 2019)، ذخایر کربن خاک (Gomes *et al.*, 2019) و بافت خاک (Malone & Searle, 2021) افزایش یافته است. در همین راستا مطالعاتی توسط محققین داخلی برای توسعه‌ی کاربرد الگوریتم‌های یادگیری ماشین در نقشه‌برداری خاک‌ها صورت پذیرفت که در آن صحرایی و همکاران (Sahraei *et al.*, 2022) با استفاده از مدل‌های ماشین‌بردار پشتیبان و جنگل تصادفی اقدام به تهیه‌ی نقشه‌های رقومی اجزای بافت خاک در بخشی از اراضی دشت خوزستان نمودند و نتایج ایشان بیانگر کارایی بالاتر مدل جنگل تصادفی نسبت به ماشین‌بردار پشتیبان بود. همچنین در مطالعه دیگری که موسوی و همکاران (Mousavi *et al.*, 2021) در ارزیابی کارایی مدل‌های یادگیری ماشین در پیش‌بینی وضعیت شوری خاک در اعماق سطحی و زیرسطحی پرداختند و نتیجه گرفتند که دو مدل کیوبیست و جنگل تصادفی نسبت به سایر مدل‌های یادگیری ماشین مورد استفاده از کارایی بالاتری برخوردار بودند بنابراین باتوجه به پتانسیل بالای الگوریتم‌های یادگیری ماشین در مدلسازی و تفسیر روابط خاک_زمین‌نما به‌عنوان ابزاری کارا در سایر دشت‌های کشاورزی کشور نیز نیاز به ارزیابی کارایی آن‌ها احساس می‌شود (Mousavi *et al.*, 2021).

از آنجایی که محصولات زراعی گندم، جو، کلزا و یونجه به‌عنوان

1- Lessivage
2- Digital Soil Mapping
3- scorpan
4- isotropic
5- Machine Learning

اصلی‌ترین محصولات در الگوی کشت این دشت می‌باشند، بنابراین به عنوان یک منطقه استراتژیک در تولید محصولات کشاورزی لحاظ می‌شود و سالانه به‌صورت پیوسته تحت کشت می‌باشد دریافت اطلاعات دقیق از وضعیت پراکنش مکانی ویژگی‌های کلیدی و تأثیرگذار بر روی کیفیت و مدیریت خاک به‌ویژه بافت خاک بسیار مهم و ضروری می‌باشد. همچنین به‌دلیل نبود نقشه‌های خاک که بیانگر وضعیت تغییرات اجزای بافت خاک در منطقه می‌باشند بنابراین پژوهش حاضر با هدف افزایش اطلاعات در مورد ۱. تهیه نقشه‌های پیش‌بینی مکانی PSD خاک ۲. مقایسه مدل‌های یادگیری ماشین RF و SVR با مدل زمین آماری کوکریجینگ و ۳. شناسایی مهمترین متغیرهای پیش‌ران PSD در دو عمق سطحی (۳۰ سانتی‌متر) و زیر سطحی (۶۰-۳۰ سانتی‌متر) در بخشی از اراضی دشت مرودشت استان فارس انجام شد.

## مواد و روش‌ها

### منطقه مورد مطالعه

در پژوهش حاضر منطقه مورد مطالعه واقع در شهرستان مرودشت در فاصله‌ی ۴۰ کیلومتری شمال شهر شیراز مرکز استان فارس قرار دارد، که دارای مساحت حدود ۵۰ هزار هکتار بخشی از منطقه دشت مرودشت می‌باشد و در موقعیت طول‌های جغرافیایی "۳۵,۸۲ ´۴۱ °۵۲ تا "۱,۰۷ ´۵۷ °۵۲ شرقی و عرض‌های جغرافیایی "۳۵,۰۲ ´۴۸ °۲۹ تا "۱۴,۷۲ ´۲ °۳۰ شمالی واقع شده است (شکل ۱-پ). متوسط ارتفاع از سطح دریای آزاد نیز ۱۶۰۰ متر می‌باشد. براساس طبقه‌بندی ماهلر[1] (Mahler, 1970) منطقه مطالعاتی دارای سه واحد فیزیوگرافی اصلی شامل کوهستان، دشت‌های دامنه‌ای و دشت‌های رسوبی بوده که براساس طبقه‌بندی ارائه شده توسط زینک (Zinck *et al.*, 2015) از نظر ژئومورفولوژی در سه واحد زمین‌نمای اصلی شامل کوهستان، دشت‌های دامنه‌ای و دشت قرار می‌گیرد. برای تعیین اقلیم منطقه مطالعاتی از داده‌های دوره آماری ۳۰ ساله هواشناسی (۱۴۰۰-۱۳۷۱) توسط ایستگاه سد درودزن (سازمان هواشناسی کشور، ۱۳۹۸) استفاده شد. میانگین بارندگی سالانه برای این منطقه ۲۸۷/۶۳ میلی‌متر و میانگین دمای سالانه نیز ۱۷/۸۰ درجه سلسیوس می‌باشد. رژیم‌های رطوبتی و حرارتی خاک‌های منطقه مطالعاتی به‌ترتیب زریک و ترمیک است و این خاک‌ها در دوره رده اصلی انتی‌سول و اینسپتی‌سول قرار دارند. مواد مادری اصلی خاک در منطقه مورد مطالعه رسوبات آبرفتی، سنگ آهک، شیل و مارن خاکستری است که قدمت آن‌ها به دوره کواترنر می‌رسد. اطلاعات زمین‌شناسی از نقشه‌های زمین شناسی ۱:۱۰۰۰۰۰ رقومی شده ایران جمع آوری شده است (Geology, 2010).

### نمونه‌برداری و مطالعات آزمایشگاهی

پس از تعیین منطقه مورد مطالعه، موقعیت ۲۰۰ نقطه نمونه‌برداری در منطقه توسط نرم‌افزار R نسخه‌ی ۴,۱,۳ با استفاده از روش ابرمکعب لاتین مشروط و دسته‌ای از متغیرهای محیطی مستخرج از مدل رقومی ارتفاع و داده‌های سنجش از دور (عمق دره، شاخص بافت پستی و بلندی، شاخص خیسی توپوگرافی، شاخص رس و شاخص تفاضلی پوشش گیاهی نرمال شده) تعیین گردید (Minasny *et al.*, 2006). به‌عبارت بهتر، برای مدلسازی ویژگی‌های خاک، ۲۰۰ نمونه از دو عمق صفر تا ۳۰ و ۳۰ تا ۶۰ سانتی‌متر در منطقه مطالعاتی برداشته شدند. سپس نمونه‌ها به آزمایشگاه منتقل شده، هوا خشک گردیده و از الک دو میلی‌متر عبور داده شدند. در نهایت، اجزای بافت خاک به روش هیدرومتر، اندازه‌گیری گردید.

### مراحل تحقیق

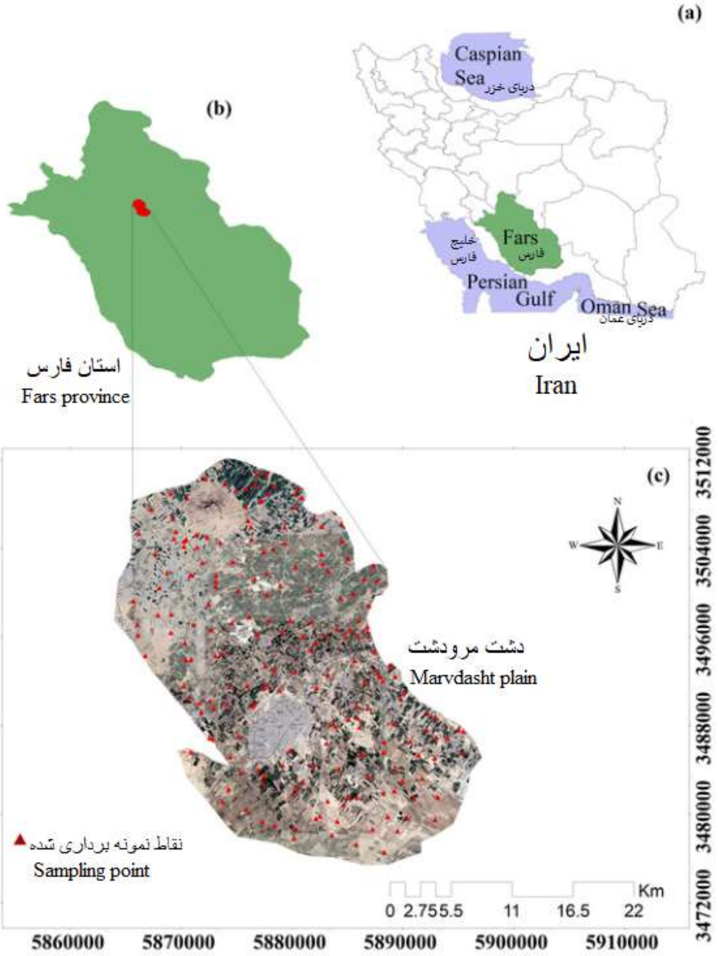
روند تهیه نقشه اجزای بافت خاک در منطقه مطالعاتی در پنج گام اصلی طراحی شد که شامل: ۱- نمونه‌برداری و جمع‌آوری اطلاعات خاک و متغیرهای محیطی از منطقه مطالعاتی، ۲- مدلسازی اجزای بافت خاک با استفاده از متغیرهای محیطی، روش‌های یادگیری ماشین و زمین آماری، ۳- اعتبارسنجی کارایی مدل‌ها در منطقه مطالعاتی، ۴- تهیه نقشه پیش‌بینی اجزای بافت خاک براساس بهترین مدل مکان مبنا، ۵- تعیین مهمترین فاکتورهای محیطی در پیش‌بینی اجزای بافت خاک.

### متغیرهای محیطی و انتخاب آن‌ها

به‌منظور شناسایی عوامل کنترل‌کننده اجزای بافت خاک در هر لایه، متغیرهای مختلفی براساس مدل اسکورپن (خاک، اقلیم، موجودات زنده، مواد مادری، پستی و بلندی، سن و موقعیت مکانی) انتخاب شدند (Mc Bratney *et al.*, 2003). متغیرهای محیطی استفاده شده در این پژوهش شامل مجموعه گسترده‌ای از نمایندگان فاکتورهای خاکساز هستند که تا حد ممکن از منابع با حداقل هزینه و با دسترسی بالا تهیه گردیدند. این متغیرها برحسب فراوانی و امکان تهیه آن‌ها شامل مشتقات اولیه و ثانویه مدل رقومی ارتفاع، شاخص‌های سنجش از دور بودند. متغیرهای محیطی پستی و بلندی از مدل رقومی ارتفاع با استفاده از نرم‌افزار SAGA GIS نسخه ۷. ۴ استخراج شدند. جهت بررسی تغییرپذیری مکانی اجزای بافت خاک از تصویر ماهواره لندست ۸ مربوط به بازه‌ی زمانی مطالعات میدانی مرداد ماه ۱۴۰۰ با قدرت تفکیک مکانی ۳۰ متر استفاده گردید همچنین قبل از استخراج شاخص‌های مورد نظر تصحیحات رادیومتری بر روی تصاویر انجام گردید. در نهایت شاخص‌های سنجش از دور نیز در محیط نرم‌افزار ENVI نسخه ۵/۳ تهیه گردیدند.

---

1- Mahler




شکل ۱- منطقه مورد مطالعاتی و محل‌های نقاط نمونه‌برداری

**Figure 1- The location of study area and sampling points**

**جدول ۱- متغیرهای محیطی قابل استفاده برای پیش‌بینی اجزای بافت خاک**

**Table 1– The applied environmental covariates for predicting soil texture particles**

| References منابع | Covariates متغیرهای محیطی | Abbreviation علامت اختصاری | Formula فرمول | Source منشاء |
|---|---|---|---|---|
| U.S. Geology Survey 2014 | Normalized different vegetation index<br>شاخص نرمال‌شده گونه‌های مختلف گیاهی | NDVI | (NIR+Red)/(NIR−Red) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Renormalized difference vegetation index<br>شاخص مجدد نرمال‌شده گونه‌های مختلف گیاهی | RDVI | (NIR+Red+L)/(NIR−Red) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Band 2 of Landsat 8<br>باند ۲ لندست ۸ | Blue<br>آبی | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Band 3 of Landsat 8<br>باند ۳ لندست ۸ | Green<br>سبز | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Band 4 of Landsat 8<br>باند ۴ لندست ۸ | Red<br>قرمز | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |

| | | | | |
|---|---|---|---|---|
| U.S. Geology Survey 2014 | Band 5 of Landsat 8<br>باند ۵ لندست ۸ | Near infrared<br>نزدیک به فروسرخ | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Band 6 of Landsat 8<br>باند ۶ لندست ۸ | Short-wave infrared-1<br>فروسرخ کوتاه موج – ۱ | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Band 7 of Landsat 8<br>باند ۷ لندست ۸ | Short-wave infrared-2<br>فروسرخ کوتاه موج – ۲ | Landsat 8 image individual band<br>تصویر لندست ۸ باند فردی | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Wetness brightness difference index<br>شاخص تفاوت روشنایی رطوبت | WBDI | (NIR−SWIR)/(NIR+SWIR) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Normalized difference salinity index<br>شاخص نرمال‌شده تفاوت شوری | NDSI | (Green−SWIR)/(Green+SWIR) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Normalized difference moisture index<br>شاخص نرمال‌شده تفاوت رطوبت | NDMI | (NIR−SWIR)/(NIR+SWIR) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Brightness Index<br>شاخص روشنایی | BI | / 3) Red+Green+Blue( | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Gypsum Index<br>شاخص گچ | GI | (SWIR1−SWIR2)/(SWIR1+SWIR2) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Clay Index<br>شاخص خاک رس | CI | (SWIR1−SWIR2)/(SWIR1+SWIR2) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Carbonate Index<br>شاخص کربنات | CI | (SWIR1−SWIR2)/(SWIR1+SWIR2) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Ratio Vegetation Index<br>شاخص نسبت گیاهی | RVI | Red / NIR | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Enhanced Vegetation Index<br>شاخص گیاهی تقویت‌شده | EVI | G× (NIR−Red) / (NIR+C1×Red−C2 ×Blue+L) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Green-Red Vegetation Index<br>شاخص گیاهی سبز–قرمز | GVI | (Green−Red)/(Green+Red) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Differenced Vegetation Index<br>شاخص تفاوت گیاهی | DVI | NIR−Red | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Infrared Percentage Vegetation Index<br>شاخص درصد گیاهی مادون قرمز | IPVI | (NIR−SWIR)/(NIR+SWIR) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Iron Oxide Ratio<br>نسبت اکسید آهن | IOR | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Soil Adjusted Vegetation Index<br>شاخص گیاهی تنظیم شده برای خاک | SVI | (NIR−Red)/ (NIR+Red+L) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Modified Soil Adjusted Vegetation Index<br>شاخص گیاهی تنظیم شده تغییر یافته برای خاک | MSAVI | )/2)2∗NIR+1−(2∗NIR+1)2−8∗(NIR−Red(( | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Perpendicular Vegetation Index<br>شاخص گیاهی عمودی | PVI | (Green−Red)/(Green+Red) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Structure Insensitive Pigment Index<br>شاخص حساس به ساختار پیگمان | SIPI | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Modified Normalized Difference Water Index<br>شاخص آب تفاوت نرمال‌شده تغییر یافته | MNDVI | (Green−SWIR)/(Green+SWIR) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Enhanced Vegetation Index<br>شاخص گیاهی تقویت‌شده | EVI | G×(NIR−Red) /(NIR+C1×Red−C2×Blue+L) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Soil Adjusted Vegetation Index<br>شاخص گیاهی تنظیم شده برای خاک | SAVI | | RS<br>سنجش از دور |

| | | | | |
|---|---|---|---|---|
| U.S. Geology Survey 2014 | Bare Soil Index<br>شاخص خاک بی‌پوشان | BSI | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Index-Based built-up Index<br>شاخص ساخت‌های مبتنی بر ایندکس | IBI | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Enhanced Vegetation Index 2<br>شاخص گیاهی تقویت‌شده۲ | EVI 2 | (NIR−Red)/ (NIR+2.4×Red+1) 2.5× | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Visible Atmospherically Resistance Index Green<br>شاخص مقاومت اتمسفری قابل رؤیت – سبز | VARIgreen | (Green−Red)/ (Green+Red−Blue) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Green Normalized Difference Vegetation Index<br>شاخص گیاهی نرمال‌شده تفاوت سبز | GNDVI | (NIR−Green)/(NIR+Green) | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Tasseled cap wetness<br>رطوبت تسلد کپ | TCW | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Tasseled cap greenness<br>سبزی تسلد کپ | TCG | | RS<br>سنجش از دور |
| U.S. Geology Survey 2014 | Normalized Ratio Vegetation Index<br>شاخص نرمال‌شده نسبت گیاهی | NRVI | NIR/Red | RS<br>سنجش از دور |
| Olaya (2004) | Wind Effect<br>تاثیر باد | WE | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Analytical Hillshading<br>تظلیل تحلیلی | AH | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Aspect<br>جهت | Aspect | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Catchment Area<br>منطقه‌ی آبگیر | CA | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Catchment Slope<br>شیب حوضه | CS | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Channel Direction<br>جهت رودخانه یا راهروی آبی | CD | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Channel Network Base Level<br>سطح پایه شبکه رودخانه | CNBL | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Channel Network Distance<br>فاصله شبکه رودخانه | CND | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Channel Network<br>شبکه رودخانه | CN | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Closed Depressions<br>تسلیمات بسته یا تسلیمات بسته‌شونده | CD | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Convergence Index<br>شاخص همگرایی | CI | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Diffuse Insolation<br>تشعشع پراکنده | Diffuse | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Direct Insolation<br>تشعشع مستقیم | DI | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Flow Directions<br>جهات جریان | FD | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Geomorphons<br>ژئومورفونز | Geomorphons | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | Landforms<br>زمین‌شکل‌ها | Landforms | | DEM<br>نقشه رقومی ارتفاع |
| Olaya (2004) | LS Factor<br>فاکتور طول و شیب | LS Factor | | DEM<br>نقشه رقومی ارتفاع |

| | | | |
|---|---|---|---|
| DEM<br>نقشه رقومی ارتفاع | MBI | Mass Balance Index<br>شاخص تعادل جرم | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | MCA | Modified Catchment Area<br>ناحیه آبگیر تغییر یافته | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | MRRTF | MRRTF | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | PC | Plan Curvature<br>انحنای مسطح | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | RSP | Relative Slope Position<br>موقعیت نسبی شیب | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SH | Slope Height<br>ارتفاع شیب | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | Slope | Slope<br>شیب | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SH | Standardized Height<br>ارتفاع استاندارد | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SPI | Stream Power Index<br>شاخص توان جریان | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SA | Surface Area<br>مساحت سطح | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | WB | Watershed Basins<br>حوضه آبخیز | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | WE | Wind Exposition<br>جهت باد | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | VD | Valley depth (m)<br>عمق دره | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | MRVBF | Multiresolution index of valley bottom flatness<br>شاخص هموارى دره با درجه وضوح بالا | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | NH | Normalized Height<br>ارتفاع نرمال‌شده | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | MS | Midslope position<br>موقعیت میانی انحدار | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | TE | Terrain surface texture<br>بافت سطح زمین | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SG | Slope gradient<br>شیب افقی | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | VDCN | Vertical Distance to Channel Network<br>فاصله عمودی تا شبکه رودخانه | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | SROI | Solar radiation or insolation<br>تابش خورشیدی | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | Convexity | Convexity<br>خمیدگی | Olaya (2004) |
| DEM<br>نقشه رقومی ارتفاع | TWI | Topographic Wetness Index<br>شاخص توپوگرافیک رطوبت | Olaya (2004) |

S= soil properties ویژگی خاک، O=organism موجودات زنده، R=relief پستی و بلندی، LAB=laboratory analysis آنالیزهای آزمایشگاهی
DEM=Digital elevation model نقشه رقومی ارتفاع، RS=Remote sensing سنجش از دور.

در مجموع، ۷۵ متغیر محیطی تهیه شد جدول ۱ که در نهایت فرمت رستری مربوط به تمام متغیرهای محیطی با اندازه سلول ۳۰ متر، متناسب با اندازه سلولی متغیرهای به‌دست آمده از تصاویر ماهواره‌ای یکسان شد. تعداد بالای متغیرهای محیطی باعث ایجاد هم‌بستگی زیاد

بین آن‌ها می‌شود که نتیجه آن، کاهش سرعت محاسبات و کاهش صحت و اعتبار مدل‌های پیش‌بینی کننده می‌باشد. برای مدل‌سازی تغییرات اجزای بافت خاک در نقشه‌برداری رقومی، علاوه بر داده‌های مشاهداتی نیاز به داده‌های کمکی مشتق شده از مدل رقومی ارتفاع و تصاویر ماهواره‌ای نیز می‌باشد. بدین منظور ۳۹ متغیر پستی و بلندی و ۳۶ متغیر سنجش از دور استخراج گردید. با توجه به حجم زیاد داده‌های موجود برای استفاده در فرآیند مدلسازی، به‌عنوان متغیرهای پیش‌بینی کننده از روش فاکتور تورم واریانس و الگوریتم باروتا در نرم افزار SPSS(V.20) برای انتخاب مناسب‌ترین مجموعه داده از میان کل متغیرهای محیطی تهیه شده و انتخاب متغیرهای محیطی بهینه برای اجرای مدل با ارائه کمترین خطا ضروری می‌باشد.

**مدل‌سازی مکانی**

**مدل‌سازی زمین آمار (سناریو ۱)**

سناریو اول صورت پذیرفته در پژوهش حاضر به بررسی روش زمین آماری برای برآورد اطلاعات اجزای بافت خاک در اعماق سطحی (Cm ۳۰-۰) و زیر سطحی (Cm۶۰-۳۰) در منطقه مورد مطالعه پرداخته شد. تمامی آنالیزهای مورد استفاده در این بخش توسط نرم‌افزار Arc GIS نسخه 10.8 بخش Geostatistical Analyst براساس روش کو-کریگینگ (Co-Kriging) صورت پذیرفت و متغیرهای محیطی منتخب برای پیش‌بینی PSD با این روش در جدول ۲ ارائه شده است. همچنین متغیرهای محیطی یا کمکی (Covariates) برای استفاده در روش کوکریجینگ به روش همبستگی پیرسون انتخاب گردیدند. قابل اشاره است که قبل از انجام برازش روش‌های زمین آماری بهتر است که از نرمال بودن توزیع داده‌ها داده‌ها اطمینان حاصل شود که در این پژوهش مقادیر چولگی بین ۱+ تا ۱- و افراشتگی بین ۳+ تا ۳- به منزله نرمال بودن توزیع داده‌ها به‌عنوان مبنا قرار گرفت.

**مدل‌سازی یادگیری ماشین (سناریو ۲)**

براساس فراوانی مدل‌های یادگیری ماشین که در مدل‌سازی ویژگی‌های خاک قابل استفاده می‌باشند، در تحقیق حاضر تلاش شده است تا کارایی این مدل‌ها در مدل‌سازی اجزای بافت خاک در دو عمق مورد نظر در منطقه مطالعاتی، مورد ارزیابی قرار گیرد و در نهایت از بین مدل‌های پیش‌بینی کننده و متغیرهای محیطی بهینه انتخاب‌شده، بهترین مدل پیش‌بینی کننده اجزای بافت خاک به همراه متغیرهای محیطی بهینه انتخاب‌شده برای به‌دست آوردن ویژگی مورد پژوهش در منطقه مطالعاتی نیز استفاده گردید. مدل جنگل تصادفی (Bi *at al.*, 2004) یکی از مدل‌های یادگیری ماشین مبتنی بر سامانه‌های درختی است که امروزه به‌عنوان یکی از پرکارترین الگوریتم‌ها در نقشه برداری رقومی خاک محسوب می‌شود (Beriman, 2001)، نحوه آموزش مدل جنگل تصادفی، به‌گونه‌ای است که در آن هر درخت به‌وسیله روش جایگزینی با تکرار از میان مجموعه‌ای از درختان به‌صورت تصادفی انتخاب و آموزش می‌بینند و زیرمجموعه‌های پیش‌بینی کننده نیز بر اساس قوانین تقسیم گره به‌صورت تصادفی انتخاب می‌گردند.

**جدول ۲- متغیرهای منتخب برای روش زمین آماری کوکریجینگ**

**Table 2– The selected covariates for Co-Kriging geostatistics method**

| Soil properties<br>ویژگی‌های مورد مطالعاتی | Depths<br>اعماق | Covariates 1<br>متغیر محیطی ۱ | Covariates 2<br>متغیر محیطی ۲ | Covariates 3<br>متغیر محیطی ۳ |
|---|---|---|---|---|
| Sand<br>شن | 0-30 | Watershed Basins<br>حوضه‌های آبخیز | MRRTF | Valley depth<br>عمق دره |
| | 30-60 | Valley depth<br>عمق دره | Watershed Basins<br>حوضه‌های آبخیز | |
| Silt<br>سیلت | 0-30 | Watershed Basins<br>حوضه‌های آبخیز | Valley depth<br>عمق دره | MRRTF |
| | 30-60 | Watershed Basins<br>حوضه‌های آبخیز | Slope<br>شیب | Valley depth<br>عمق دره |
| Clay<br>رس | 0-30 | Convergence index<br>شاخص همگرایی | Watershed Basins<br>حوضه‌های آبخیز | Valley depth<br>عمق دره |
| | 30-60 | Valley depth<br>عمق دره | Watershed Basins<br>حوضه‌های آبخیز | Slope<br>شیب |

در حین فرآیند اجرای مدل RF، تعداد درختان در هر جنگل (ntree)و تعداد متغیرهای مورد استفاده برای رشد هر درخت (mtry)، به‌عنوان مهم‌ترین پارامترهای تنظیمی مدل RF می‌باشند. در این تحقیق جهت بهینه‌سازی فاکتورهای اصلی و تنظیم‌کننده کارکرد مدل RF به‌ترتیب تعداد درختان (ntree) با فاصله هر ۵۰ عدد در دامنه ۱۰۰ تا ۱۰۰۰ درخت با فاصله ۱۰۰ عدد درخت و تعداد متغیرهای محیطی قابل استفاده برای رشد هر درخت بر اساس حداقل مقدار خطا RMSEبر اساس بسته تخصصی "random forest" در محیط نرم افزار R تعیین گردیدند (Friedman & Meulman, 2003). مدل دیگر مورد استفاده در این پژوهش رگرسیون ماشین‌بردار پشتیبان می باشد، این روش الگوریتمی است که رابطه میزان ورودی و خروجی یک سیستم را از روی نمونه‌های موجود و یا داده‌های آموزش، تخمین می‌زند (Bi *et al.*, 2004). تمامی فرآیندهای مدلسازی در محیط نرم‌افزار Rstudio 1.0.136 صورت پذیرفت.

### اعتبارسنجی نتایج مدلسازی

فرایند اعتبارسنجی مدل‌های یادگیری ماشین و زمین آماری مورد استفاده با استفاده از ۲۰ درصد کل داده‌ها صورت پذیرفت و صحت مدل‌ها با استفاده از شاخص‌های آماری ضریب تبیین ($R^2$)، ریشه نرمال شده میانگین مربعات خطا (nRMSE) و درصد اختلاف نسبی (RPD) محاسبه شد (روابط ۱، ۲ و ۳) (Rossel & McBratney, 2008).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad (۱)$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - o_i)^2}}{\bar{O}} \quad (۲)$$

$$RPD = \frac{SD}{RMSE} \times 100 \quad (۳)$$

که در روابط بالا، n تعداد نقاط نمونه‌برداری، Oi مقدار مشاهداتی، $\bar{O}$ میانگین مقادیر مشاهداتی و Pi مقدار تخمین زده شده متغیر، SD انحراف استاندارد مقادیر مشاهداتی RMSE ریشه دوم میانگین مربعات خطا.

RPD پایین‌تر نشان دهنده تطابق بهتر بین مقادیر مشاهداتی و پیش‌بینی شده است. مقادیر RPD را می‌توان برای طبقه‌بندی دقت یک روش پیش‌بینی به‌صورت زیر استفاده کرد:

میزان RPD > ۱ دقت ضعیف، میزان ۱ ≤ RPD> ۲ دقت متوسط، میزان ۲ ≤ RPD> ۳ دارای دقت رضایت بخش، میزان ۳ ≤ RPD> ۴ دارای دقت خوب و میزان RPD ≥ ۴ دارای دقت عالی می‌باشد (Sørensen *et al.*, 2004).

## نتایج و بحث

### خلاصه آماری اجزای بافت خاک

خلاصه آماری اجزای بافت خاک اندازه‌گیری شده در دو عمق سطحی و زیر سطحی منطقه مورد مطالعه در جدول ۳ ارائه شده است. کمترین میزان رس به‌ترتیب ۱۰/۲۱ و ۱۰/۴۵ و بیشترین میزان آن ۶۵/۳۲ و ۶۳/۳۶ درصد در عمق سطحی و زیر سطحی اندازه‌گیری شد. میانگین رس در تمام نمونه‌ها ۳۷/۹۱ و ۳۵/۶۱ درصد است. میانگین شن در عمق سطحی ۲۵/۶۵ و در عمق زیر سطحی ۲۶/۰۲ درصد اندازه‌گیری شد. به‌طوری‌که بیشترین میزان آن در عمق سطحی و برابر با ۶۸/۵۴ درصد و کمترین مقدار آن در عمق زیرسطحی و برابر با ۳/۷۶ درصد است. بیشترین و کمترین مقدار سیلت در عمق زیرسطحی و به‌ترتیب برابر با ۷۶/۱۳ و ۴/۹۱ با میانگین ۳۸/۳۷ درصد است. براساس طبقه‌بندی (Wilding, 1985) تغییرپذیری شن در منطقه مورد مطالعه زیاد (بیش از ۳۵ درصد) بوده و رس و سیلت دارای تغییرپذیری متوسط (۱۳–۳۵ درصد) می‌باشند. زراعت پیشه و همکاران (Zeraatpisheh *et al.*, 2020) نیز میانگین رس در استان فارس را ۳۷/۹ درصد و تغییرپذیری آن را متوسط (۲۰/۱ درصد) گزارش کردند. تشیع و همکاران (Tashayo *et al.*, 2020) نیز گزارش کردند که میانگین شن در مرودشت استان فارس بیشتر از رس و سیلت است.

### انتخاب متغیرهای کمکی محیطی

مناسب‌ترین متغیرهای کمکی با استفاده از آزمون هم خطی توسط شاخص تورم واریانس (VIF) و سپس رویکرد انتخاب متغیر باروتا انتخاب شدند و در جدول ۴ و شکل ۲ ارائه شده است. چهار متغیر عمق دره (VD)، بافت (TE)، شاخص رطوبت پستی و بلندی (TWI) و شاخص رس (CM) مربوط به پارامترهای ژئومرفومتری و متغیر شاخص تفاضلی پوشش گیاهی نرمال‌شده (NDVI) مربوط به شاخص‌های سنجش از دور می‌باشند (جدول ۴). جابجایی ذرات ریز خاک ارتباط نزدیکی با توپوگرافی منطقه مورد مطالعه دارد ( Swain *et al.*, 2021). مصلح و همکاران (Mosleh *et al.*, 2016) و موسوی و همکاران (Mousavi *et al.*, 2019) بیان نمودند که پارامترهای استخراج‌شده از مدل رقومی ارتفاع می‌توانند به‌عنوان متغیرهای محیطی مناسب در مدل‌سازی کلاس و خصوصیات خاک محسوب شوند.

**جدول ۳- خلاصه آماری اجزای بافت خاک در منطقه مطالعاتی در دو عمق سطحی و زیرسطحی**

**Table 3- Statistical summary of Soil Texture Particles in study area at surface and subsurface depths**

| اجزای بافت خاک<br>Soil texture | اعماق<br>Depths | میانگین<br>Mean | حداقل<br>Min | حداکثر<br>Max | انحراف معیار<br>SD | ضریب تغییرات%<br>CV |
|---|---|---|---|---|---|---|
| Sand | 0-30 | 25.6 | 8.50 | 68.54 | 11.2 | 0.45 |
| شن | 30-60 | 26.0 | 3.76 | 66.06 | 12.58 | 0.48 |
| Silt | 0-30 | 36.4 | 10.13 | 70.66 | 10.7 | 0.29 |
| سیلت | 30-60 | 38.3 | 4.91 | 76.13 | 10.69 | 0.30 |
| Clay | 0-30 | 37.9 | 10.21 | 65.32 | 9.78 | 0.25 |
| رس | 30-60 | 35.6 | 10.45 | 63.36 | 9.90 | 0.25 |

CV: Coefficient of variation, SD: Standard deviation, Max: Maximum, Min: Minimum

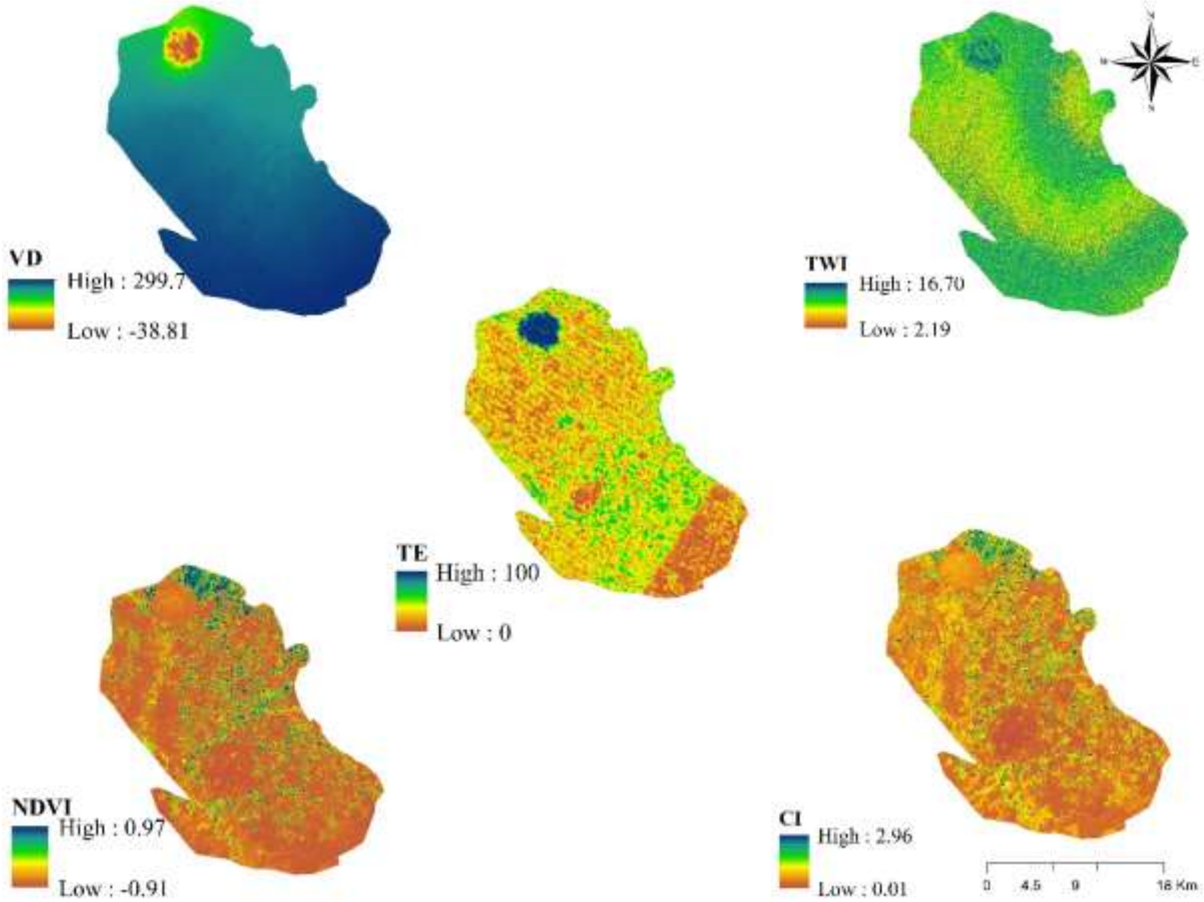
**جدول ٤- متغیرهای کمکی محیطی مورد استفاده در پیش‌بینی مکانی اجزای بافت خاک**

**Table 4- Environmental covariates used in Soil Texture Particles spatial prediction**

| فاکتور خاک‌ساز (scorpan)<br>Soil Forming Factor | نماد<br>Symbol | متغیرکمکی محیطی<br>Environmental variable | منبع<br>Source |
|---|---|---|---|
| پستی و بلندی<br>Topography (r) | VD | عمق دره<br>Valley depth | مدل رقومی ارتفاع<br>Digital elevation model |
| پستی و بلندی<br>Topography (r) | TE | شاخص بافت پستی و بلندی<br>Terrain surface texture | |
| پستی و بلندی<br>Topography (r) | TWI | شاخص خیسی پستی و بلندی<br>Topographic wetness Index | |
| پستی و بلندی<br>Topography (r) | CI | شاخص رس<br>Clay Index | |
| زیستی<br>Organism (O) | NDVI | شاخص تفاضلی پوشش گیاهی نرمال‌شده<br>Normalized differences vegetation index | داده های سنجش از دور<br>Remote sensing |

r و o به‌ترتیب نشانگر فاکتورهای پستی و بلندی و زیستی در مدل scorpan هستند.

r, p represents topography and organism factors in the scorpan model, respectively.




**شکل ۲- متغیرهای کمکی محیطی مورد استفاده در پیش‌بینی مکانی اجزای بافت خاک**
**Figure 2- Environmental covariates used in Soil Texture Particles spatial prediction**

همچنین مطالعات مختلف گزارش نمودند که داده‌های سنجش از دور در سطح جهانی و محلی در مطالعات نقشه‌برداری رقومی خاک موفق بوده است (Hengl *et al*., 2017). چن و همکاران (Chen *et al*., 2022) نشان دادند که استفاده از مدل باروتا برای انتخاب متغیرهای محیطی کمکی نسبت به سایر مدل‌های انتخاب متغیر منجر به خطا (RMSE) کمتر برای پیش‌بینی کربن‌آلی خاک شد.

### مقایسه ارزیابی عملکرد مدل‌ها

عملکرد مدل‌های یادگیری ماشین جنگل تصادفی (RF) و رگرسیون بردار پشتیبان (SVR) و مدل زمین‌آماری کوکریجینگ (Co-kriging) با استفاده از سه شاخص آماری $R^2$، nRMSE و RPD محاسبه شد (جدول ۵). مقادیر $R^2$ برای پیش‌بینی شن در عمق سطحی توسط مدل RF، SVR و Co-Kriging به‌ترتیب ۰/۵۴، ۰/۱۶ و ۰/۱۶ و برای عمق زیرسطحی به‌ترتیب ۰/۳۶، ۰/۱۵ و ۰/۰۶ درصد محاسبه شد. برای پیش‌بینی سیلت میزان $R^2$ برای RF به‌ترتیب برابر با ۰/۴۸ و ۰/۶۴، برای SVR به‌ترتیب برابر با ۰/۱۰ و ۰/۰۶، و برای Co-Kriging به‌ترتیب برابر با ۰/۰۰۲ و ۰/۰۱ در دو عمق سطحی و زیر سطحی به دست آمد. برای پیش‌بینی رس نیز مدل RF با $R^2$ ۰/۵۲ و ۰/۴۹ بهترین عملکرد را دارد. مقایسه نتایج nRMSE و RPD نشان داد که مدل RF برای پیش‌بینی رس (nRMSE به‌ترتیب برابر با ۰/۱۶ و ۰/۲۷ و $1 \leq RPD < 2$) دارای عملکرد متوسط و دقت متوسط مشابه با پیش‌بینی سیلت (nRMSE به‌ترتیب برابر با ۰/۱۹ و ۰/۲۴ و $1 \leq RPD < 2$) بود و دقت رضایت بخش و شن (nRMSE به‌ترتیب ۰/۳۲ و ۰/۴۲ و $1 \leq RPD < 2$) نیز عملکردی شبیه دو ویژگی دیگر از خود نشان داد (Wallach *et al*., 2006; Sørensen *et al*., 2004) که می‌تواند به علت تغییرپذیری کمتر رس نسبت به سیلت و شن باشد (جدول ۵). به‌طور کلی پیش‌بینی شن و رس در عمق سطحی دارای دقت بالاتر نسبت به عمق زیرسطحی است که با نتایج لوئیزو و همکاران (Loiseau *et al*., 2019) مطابقت داشت.

**جدول ۵- نتایج اعتبارسنجی مدل زمین آماری و مدل‌های یادگیری ماشین برای اجزای بافت خاک در عمق سطحی و زیرسطحی در منطقه مورد مطالعه**

**Table 5- Validation results of the geostatistics model and machine learning models for Soil Texture Particles at surface and subsurface depth in the study areas**

| ویژگی Property | بهترین مدل پیش‌بینی کننده The best predictive model | عمق Depth (cm) | $R^2$ | nRMSE | PRD |
|---|---|---|---|---|---|
| نتایج اعتبارسنجی در منطقه مطالعاتی | | | | | |
| Sand شن | RF | 0-30 | 0.54 | 0.32 | 1.35 |
| | | 30-60 | 0.36 | 0.42 | 1.15 |
| | SVR | 0-30 | 0.16 | 0.62 | 0.70 |
| | | 30-60 | 0.15 | 0.86 | 0.56 |
| | Co-kriging | 0-30 | 0.16 | 0.97 | 0.45 |
| | | 30-60 | 0.06 | 0.87 | 0.55 |
| Silt سیلت | RF | 0-30 | 0.48 | 0.19 | 1.47 |
| | | 30-60 | 0.64 | 0.24 | 1.21 |
| | SVR | 0-30 | 0.10 | 0.32 | 0.90 |
| | | 30-60 | 0.06 | 0.34 | 0.83 |
| | Co-kriging | 0-30 | 0.002 | 0.86 | 0.34 |
| | | 30-60 | 0.01 | 0.90 | 0.31 |
| Clay رس | RF | 0-30 | 0.52 | 0.16 | 1.67 |
| | | 30-60 | 0.49 | 0.27 | 1.01 |
| | SVR | 0-30 | 0.08 | 0.27 | 0.95 |
| | | 30-60 | 0.11 | 0.27 | 1.03 |
| | Co-kriging | 0-30 | 0.04 | 0.46 | 0.56 |
| | | 30-60 | 0.07 | 0.66 | 0.42 |

شهریاری و همکاران (Shahriari *et al.*, 2019) نیز اعلام کردند که مدل RF نسبت به مدل‌های زمین‌آماری دارای عملکرد بهتری برای پیش‌بینی شن، سیلت و رس است. مدل جنگل تصادفی می‌تواند روابط غیرخطی بین متغیرهای محیطی و ویژگی‌های خاک را به خوبی در حین فرآیند مدلسازی و پیش‌بینی لحاظ نماید همچنین قادر است تا با مدیریت داده‌های پرت از بروز بیش برازش و کم برازش جلوگیری نماید (Azizi *et al.*, 2023; Khosravani *et al.*, 2023a). نتایج موسوی و همکاران (Mousavi *et al.*, 2020) نیز بیانگر دقت بالای مدل RF برای پیش‌بینی ویژگی‌های سطحی خاک شامل کربن‌آلی، کربنات کلسیم معادل، رس و سیلت نسبت به روش کریجینگ معمولی است. نتایج صحرائی و همکاران (Sahraei *et al.*, 2022) نیز نشان داد که مدل RF نسبت به مدل ماشین بردار پشتیبان (SVM)[1] دارای عملکر بهتر برای پیش‌بینی اجزای بافت خاک است. آنها اعلام کردند که مدل SVM به لحاظ تغییرپذیری نسبت به مدل‌های پیچیده‌تر مانند RF، عدم قطعیت بیشتری را نشان می‌دهند، بنابراین مدل RF نسبت به مدل SVM دقت بالاتری را برای پیش‌بینی ذرات خاک دارا می‌باشد. نتایج ژانگ و همکاران (Zhang *et al.*, 2023) نیز نشان داد که میزان رس خاک و تنوع آن با عمق خاک افزایش می یابد. به‌طوری‌که میانگین خطای مطلق (MAE)[2]، ریشه میانگین مربعات خطا (RMSE)، ضریب تعیین ($R^2$) و ضریب همبستگی تطابق لین (CCC)[3] به‌ترتیب از ۶/۹۵ تا ۸/۷۴، ۸/۹ تا ۱۱/۳، ۴۲ تا ۴۶ و ۵۶ تا ۵۹ درصد متغیر بود. همچنین تومساواش و همکاران (Tu¨msavas *et al.*, 2019) RPD ۳/۲۵ و ۳/۵۱ را برای پیش‌بینی شن و سیلت گزارش کردند.

## اهمیت متغیرهای محیطی

تجزیه و تحلیل متغیرهای مهم در پیش‌بینی مدل‌های ML می تواند بینش‌هایی در مورد ماهیت مجموعه داده مورد استفاده ارائه دهد و در عین حال به افزایش کارایی پیش‌بینی‌های ML نیز کمک کند. بررسی ویژگی‌ها همچنین امکان تخصیص امتیازها به چندین ورودی را فراهم می‌کند تا مهم‌ترین تغییرات در مدل‌سازی را شناسایی کند. اهمیت متغیرهای محیطی استفاده شده برای پیش‌بینی تغییرات مکانی ذرات خاک در اعماق سطحی و زیرسطحی توسط مدل RF در شکل ۳ نشان داده شده است. بر این اساس VD با ۶۰ و ۶۵ درصد برای پیش‌بینی شن در عمق سطحی و زیر سطحی، ۷۰ درصد برای پیش‌بینی سیلت در عمق سطحی و ۷۰ و ۶۵ درصد برای پیش‌بینی رس در هر دو عمق سطحی و زیر سطحی دارای بیشترین اهمیت است. تنها برای پیش‌بینی سیلت در عمق زیرسطحی متغیرهای TE و TWI دارای اهمیت بیشتری نسبت به VD می‌باشند (شکل ۳). این نتایج نشان

1- Support Vector Machine
2- Mean Absolute Error
3- Lin’s Concordance Correlation Coefficient

می‌دهد که متغیرهای توپوگرافی بر تغییرات مکانی ذرات خاک مؤثر می‌باشند. این یافته با نتایج تقی زاده مهرجردی و همکاران (Taghizadeh-Mehrjardi *et al.*, 2016; Maleki *et al.*, 2023; Khosravani *et al.*, 2023b) و پهلوان راد و اکبری مقدم، (Pahlavan-Rad & Akbarimoghaddam, 2018) منطبق می باشد. توپوگرافی بر توزیع اندازه ذرات خاک عمدتا با کنترل جریان آب و حمل و نقل ذرات خاک تأثیر می‌گذارد (Mahmoudzadeh et *al.*, 2020). ژانگ و همکاران (Zhang *et al.*, 2019) دریافتند که خاک سطحی به شدت در معرض کنترل‌های توپوگرافی قرار دارد. در مطالعه حاضر VD مهمترین عامل توپوگرافی است که توزیع ذرات خاک را تحت تأثیر قرار داده است. VD بیانگر فاصله عمودی در سطح پایه شبکه کانال است، بنابراین عامل مهمی برای تعیین کمیت کنترل توپوگرافی در فرایندهای هیدرولوژیکی است (Lee *et al.*, 2020). پهلوان راد و اکبری مقدم، (Pahlavan-Rad & Akbarimoghaddam, 2018) فاصله تا شبکه آبراهه[1] را مهمترین متغیر در پیش‌بینی رس شناسایی کردند. جان و همکاران (John *et al.*, 2020) متغیرهای محیطی TWI و حوزه آبریز را مهمترین متغیرهای برای پیش‌بینی ویژگی‌های خاک معرفی کردند. دهروماراجان و هگده (Dharumarajan & Hegde, 2022) نیز شاخص همواری کف دره با درجه تفکیک بالا[2] (MrVBF) را به‌عنوان مهمترین متغیر در پیش‌بینی توزیع مکانی ذرات خاک معرفی کردند. کایا و باشاقیت (Kaya & Başayiğit, 2021) بیان کردند که مدل رقومی ارتفاع و مشتقات آن مهمترین متغیرهای محیطی در پیش‌بینی توزیع ذرات خاک است.

## توزیع مکانی ذرات خاک

با توجه به اینکه مدل RF نسبت به دو مدل SVM و Cokriging از دقت بالاتری برای پیش‌بینی اجزای بافت خاک برخوردار بود بنابراین نقشه‌های توزیع مکانی ذرات خاک توسط مدل RF در شکل ۴ نشان داده شده است. براساس شکل ۴ حداقل و حداکثر رس در عمق سطحی به ترتیب ۲۰/۹۳ و ۵۲/۵۷ و در عمق زیر سطحی به‌ترتیب ۲۰/۴۲ و۵۴/۹۴ (شکل ۴ پ و ج) درصد مشاهده گردید. کمترین مقدار رس در هر دو عمق سطحی و زیر سطحی در قسمت شمال منطقه مورد مطالعه پیش‌بینی شد. در حالی‌که بیشترین مقدار آن در بخش غرب، مرکز تا جنوب وجود دارد.

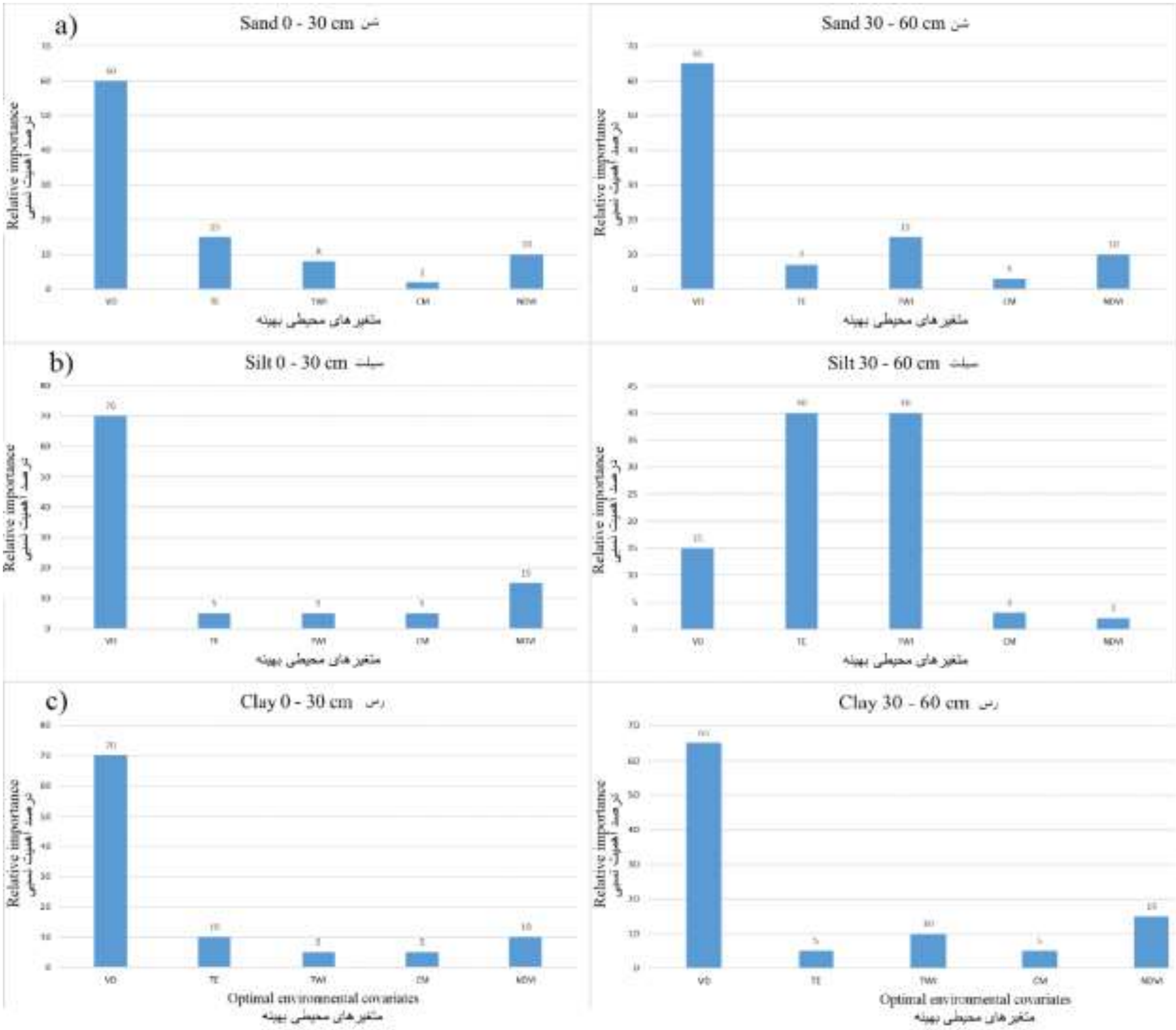
چنین به‌نظر می‌رسد که پایین بودن مقدار رس در بخش‌های شمالی وابسته به تأثیر فرآیندهای فرسایش و رسوب به‌همراه اثر توپوگرافی می‌باشد، به‌طوری‌که مناطق مرتفع و با ارتفاع بیشتر در بخش شمالی بوده که تلفیق اثر توپوگرافی و انتقال املاح منجر به کاهش مقدار رس در مناطق شمالی و افزایش آن در مناطق جنوبی شده است. مشاهده‌ی این الگوی پراکنش برای رس با نمودارهای اهمیت نسبی آن (شکل ۳ ب و شکل ۲) همراستا هست به‌طوری‌که شاخص VD با بالاترین میزان اهمیت بیشترین نقش را در تعیین الگوی پراکنش مکانی رس ایفا نموده است، نتایج مشابهی نیز توسط (Ließ *et al.*, 2012) گزارش شد و بیان نمودند که مقدار رس در زمین‌نما با افزایش ارتفاع و شیب کاهش و با انحنای شیب افزایش می‌یابد.

الگوی تغییرات مکانی شن (عمق سطحی و زیرسطحی) و سیلت (عمق سطحی) براساس شکل ۴ روند مشابهی را نشان داد. بیشترین و کمترین مقدار شن پیش‌بینی شده در عمق سطحی ۱۵/۱۴ و ۵۴/۲۴ و زیر سطحی به‌ترتیب برابر با ۱۴/۴۶ و ۵۵/۱۷ (شکل ۴ ب و ث) و برای سیلت در عمق سطحی برابر ۱۷/۲۳ و ۷۳/۶۰ (شکل ۴ الف) درصد مشاهده گردید. برخلاف رس بیشترین مقدار شن در هر دو عمق در قسمت شمالی و مرتفع‌ترین قسمت منطقه مورد مطالعه مشاهده گردید و کمترین مقدار آن در نواحی پست منطقه مورد مطالعه مشاهده گردید. این نتایج در حالی است که طبق (شکل ۴ ت) روند تغییرات سیلت در عمق زیرسطحی به الگوی رس مشابه‌تر می‌باشد. مناطق پست و همچنین دشت های رسوبی در قسمت مرکزی منطقه مورد مطالعه حاوی مقادیر زیادی سیلت بودند. بنابراین می‌توان بیان نمود که این تغییرات در دو عمق سطحی و زیرسطحی برای اجزای بافت خاک تا حد زیادی متاثر از توالی فرآیندهای فرسایش و رسوب بوده که توزیع نامنظمی از اجزای بافت را از سطح به عمق نشان می‌دهد. در همین راستا ویلکی و همکاران (Wilcke *et al.*, 2008) ارتباط قوی بین اجزای بافت خاک و ارتفاع در منطقه‌ی مطالعاتی خود توصیف نمود و بیان نمود که یک همبستگی مثبت بین پارامترهای وابسته به ارتفاع با شن و متعاقبا همبستگی منفی با مقدار رس وجود دارد. همچنین نتایج مشابهی توسط سایر محققین (Gessler *et al.*, 2000; Ziadat, 2005; Riza *et al.*, 2021; Mousavi *et al.*, 2023) در مورد ارتباط بین اجزای بافت خاک و تغییرات ارتفاع مشاهده شده است. فرآیندهای توپوگرافی، با کنترل فرآیندهای فرسایش و رسوب بر روی مقادیر پراکنش مکانی اجزا افت خاک مؤثر هستند و همچنین در مناطق مرتفع (قله و شانه شیب) بیشترین مقدار شن و در مناطق پای شیب بیشترین مقدار رس وجود دارد.

این نقشه‌ها می‌توانند به‌عنوان متغیرهای کمکی در سایر برنامه ریزی‌های مربوط به مدیریت اراضی از قبیل تعیین آب قابل استفاده و حاصلخیزی خاک مورد استفاده قرار گیرند (Ostovari *et al.*, 2021).

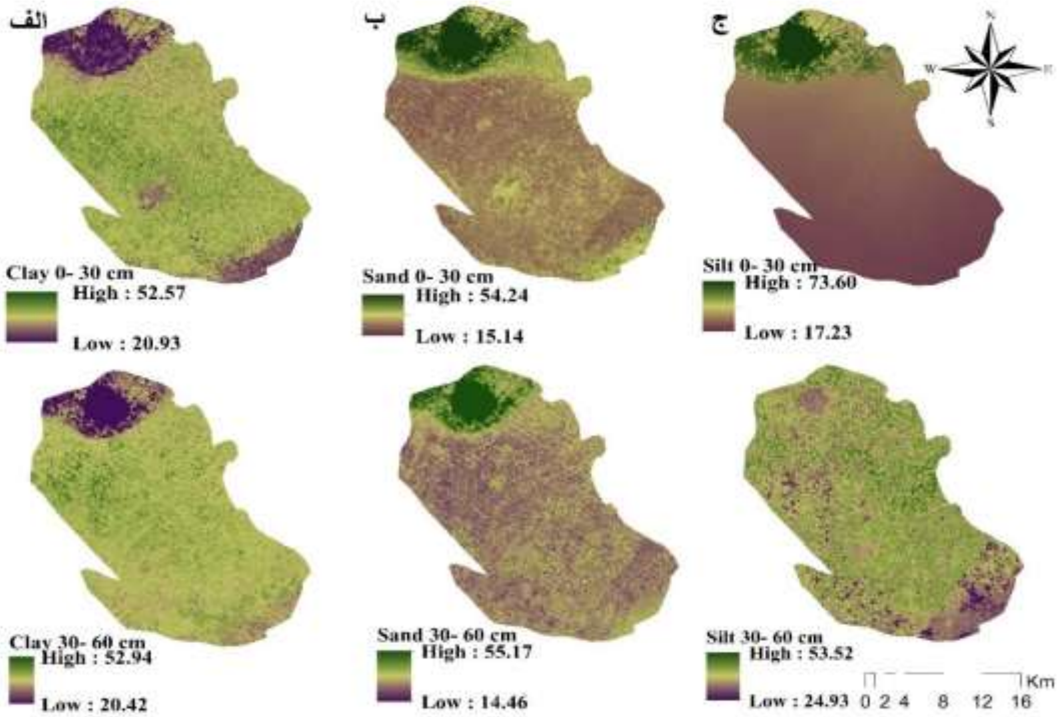
1- Channel Network Distance

2- Multi Resolution Index of Valley Bottom Flatness




**شکل ۳- نمودار اهمیت متغیرهای محیطی برای پیش‌بینی الف. شن، ب. سیلت و پ. رس در عمق سطحی و زیر سطحی**

**Figure 3- Diagram of importance of environmental variables for predicting a. sand, b. silt and c. clay at surface and subsurface depth**




**شکل ٤- نقشه توزیع الف) سیلت سطحی، ب) شن سطحی، پ) رس سطحی، ت) سیلت زیرسطحی ث) شن زیرسطحی، ج) رس زیرسطحی پیش‌بینی شده توسط مدل RF**

**Figure 4- The distribution map of Sand, Silt and Clay predicted in the surface and subsurface depth by the RF model (From top right to bottom left).**

## نتیجه‌گیری

پژوهش حاضر با هدف کلی تهیه‌ی نقشه‌های اجزای بافت خاک (رس، سیلت و شن) و شناسایی مهمترین فاکتورهای پیش‌ران آن‌ها در بخشی از اراضی دشت مروشت با کمک مدل‌های یادگیری ماشین و زمین آماری انجام گردید. مدل RF نسبت به مدل SVR و Co-Kriging در پیش‌بینی شن، سیلت و رس در هر دو عمق سطحی (۳۰ سانتی‌متر) و زیر سطحی (۶۰–۳۰ سانتی‌متر) عملکرد بهتری داشت. با توجه به پاسخ طیفی خاک که توسط عوامل متعددی (مانند رطوبت خاک، مواد آلی، زبری سطح، اثرات جوی، اثرات ساختاری، مدیریت خاک و پوشش رویشی) کنترل می‌شود، داده‌های مشتق شده از مدل رقومی ارتفاع همراه با داده‌های سنجش از دور در عملکرد مدل برای پیش‌بینی با دقت و صحت بیشتر مؤثر بود. به‌طوری‌که متغیر VD به عنوان مهمترین متغیر در پیش‌بینی توزیع مکانی ذرات خاک در منطقه مورد مطالعه شناسایی شد. برای افزایش دقت پیش‌بینی، می‌توان در مطالعات آتی از سایر منابع متغیرهای کمکی مانند متغیرهای آب و هوایی و داده‌های راداری استفاده کرد. همچنین رویکرد مورد استفاده در این تحقیق می‌تواند به‌عنوان راهنمایی برای تهیه‌ی نقشه‌های مدیریت پذیر خاک مورد استفاده قرار گیرد.

## منابع

1. Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A.C., Morgan, C.L., & Mulder, V.L. (2020). Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional*, *20*, e00255. https://doi.org/10.1016/j.geodrs.2020.e00255
2. Azizi, K., Garosi, Y., Ayoubi, S., & Tajik, S. (2023). Integration of Sentinel-1/2 and topographic attributes to predict the spatial distribution of soil texture fractions in some agricultural soils of western Iran. *Soil and Tillage Research*, *229*, 105681. https://doi.org/10.1016/j.still.2023.105681
3. Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32. https://doi.org/10.1023/A:1010933404324
4. Bi, D., Li, Y.F., Tso, S.K., & Wang, G.L. (2004). Friction modeling and compensation for haptic display based on support vector machine. *IEEE Transactions on Industrial Electronics*, *51*(2), 491-500. https://doi.org/10.1109/TIE.2004.825277

5. Dharumarajan, S., & Hegde, R. (2022). Digital mapping of soil texture classes using Random Forest classification algorithm. *Soil Use and Management*, *38*(1), 135-149. https://doi.org/10.1111/sum.12668
6. de Jesus Duarte, S., Glaser, B., & Pellegrino Cerri, C.E. (2019). Effect of biochar particle size on physical, hydrological and chemical properties of loamy and sandy tropical soils. *Agronomy*, *9*(4), 165. https://doi.org/10.3390/agronomy9040165
7. Chen, T.L., Shi, Z.L., Wen, A.B., Yan, D.C., Guo, J., Chen, J.C., & Chen, R.Y. (2021). Multifractal characteristics and spatial variability of soil particle-size distribution in different land use patterns in a small catchment of the Three Gorges Reservoir Region, China. *Journal of Mountain Science*, *18*(1), 111-125. https://doi.org/10.1007/s11629-020-6112-5
8. Chen, Y., Ma, L., Yu, D., Zhang, H., Feng, K., Wang, X., & Song, J. (2022). Comparison of feature selection methods for mapping soil organic matter in subtropical restored forests. *Ecological Indicators*, *135*, 108545. https://doi.org/10.1016/j.ecolind.2022.108545
9. Faé, G.S., Montes, F., Bazilevskaya, E., Añó, R.M., & Kemanian, A.R. (2019). Making soil particle size analysis by laser diffraction compatible with standard soil texture determination methods. *Soil Science Society of America Journal*, *83*(4), 1244-1252. http://doi.org/10.2136/sssaj2018.10.0385
10. Friedman, J.H., & Meulman, J.J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine*, *22*(9), 1365-1381. https://doi.org/10.1002/sim.1501
11. Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., & Holmes, K. (2000). Modeling soil–landscape and ecosystem properties using terrain attributes. *Soil Science Society of America Journal*, *64*(6), 2046-2056. https://doi.org/10.2136/sssaj2000.6462046x
12. Geology.com/news/2010/freelansatimages-from-USGS-2. http://glovis.usgs.gov.
13. Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G., & Fernandes Filho, E.I. (2019). Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, *340*, 337-350. https://doi.org/10.1016/j.geoderma.2019.01.007
14. Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748
15. Hossain, M.S., Rahman, G.M., Alam, M.S., Rahman, M.M., Solaiman, A.R.M., & Mia, M.B. (2018). Modelling of soil texture and its verification with related soil properties. *Soil Research*, *56*(4), 421-428. https://doi.org/10.1071/sr17252
16. Jenny, H. (1994). *Factors of soil formation: a system of quantitative pedology*. Courier Corporation.
17. John, K., Abraham Isong, I., Michael Kebonye, N., Okon Ayito, E., Chapman Agyeman, P., & Marcus Afu, S. (2020). Using machine learning algorithms to estimate soil organic carbon variability with environmental variables and soil nutrient indicators in an alluvial soil. *Land*, *9*(12), 487. https://doi.org/10.3390/land9120487
18. Kaya, F., & Başayiğit, L. (2022). Spatial prediction and digital mapping of soil texture classes in a Floodplain using multinomial Logistic regression. In *Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation: Proceedings of the INFUS 2021 Conference, held August 24-26, 2021. Volume 2* (pp. 463-473). Springer International Publishing. https://doi.org/10.1007/978-3-030-85577-2_55.
19. Khosravani, P., Baghernejad, M., Moosavi, A.A., & FallahShamsi, S.R. (2023). Digital mapping to extrapolate the selected soil fertility attributes in calcareous soils of a semiarid region in Iran. *Journal of Soils and Sediments*, *23*(11), 4032-4054. https://doi.org/10.1007/s11368-023-03548-1
20. Khosravani, P., Baghernejad, M., Moosavi, A.A., & Rezaei, M. (2023). Digital mapping and spatial modeling of some soil physical and mechanical properties in a semi-arid region of Iran. *Environmental Monitoring and Assessment*, *195*(11), 1367. https://doi.org/10.1007/s10661-023-11980-6
21. Lee, S., Baek, W.K., Jung, H.S., & Lee, S. (2020). Susceptibility mapping on urban landslides using deep learning approaches in Mt. Umyeon. *Applied Sciences*, *10*(22), 8189. https://doi.org/10.3390/app10228189
22. Loiseau, T., Chen, S., Mulder, V.L., Dobarco, M.R., Richer-de-Forges, A.C., Lehmann, S., ... & Arrouays, D. (2019). Satellite data integration for soil clay content modelling at a national scale. *International Journal of Applied Earth Observation and Geoinformation*, *82*, 101905. https://doi.org/10.1016/j.jag.2019.101905
23. Lucas, M., Schlüter, S., Vogel, H.J., & Vetterlein, D. (2019). Soil structure formation along an agricultural chronosequence. *Geoderma*, *350*, 61-72. https://doi.org/10.1016/j.geoderma.2019.04.041
24. Ma, Y., Minasny, B., Malone, B.P., & Mcbratney, A.B. (2019). Pedology and digital soil mapping (DSM). *European Journal of Soil Science*, *70*(2), 216-235. https://doi.org/10.1111/ejss.12790.
25. Mahler, P.J. (1970). Manual of Multipurpose Land Classification. Report no. 212. *Soil and Water Research Institute*, Tehran. Iran. (In Persian)
26. Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrjardi, R., Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, *21*, e00260. https://doi.org/10.1016/j.geodrs.2020.e00260

27. Maleki, S., Karimi, A., Mousavi, A., Kerry, R., & Taghizadeh-Mehrjardi, R. (2023). Delineation of soil management zone maps at the regional scale using machine learning. *Agronomy*, *13*(2), 445. https://doi.org/10.3390/agronomy 13020445
28. Malone, B., & Searle, R. (2021). Updating the Australian digital soil texture mapping (Part 1*): re-calibration of field soil texture class centroids and description of a field soil texture conversion algorithm. *Soil Research*, *59*(5), 419-434. https://doi.org/10.1071/SR20283
29. McBratney, A.B., Santos, M.M., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, *117*(1-2), 3-52. https://doi.org/10.1016/S0016-7061(03)00223-4
30. Minasny, B., & McBratney, A.B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & geosciences*, *32*(9), 1378-1388.https://doi.org/10.1016/j.cageo.2005.12.009
31. Mosleh, Z., Salehi, M. H., Jafari, A., Borujeni, I.E., & Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring and Assessment*, *188*, 1-13. https://doi.org/10.1007/s10661-016-5204-8
32. Mousavi, S.R., Sarmadian, F., Omid, M., & Bogaert, P. (2021). Digital modeling of three-dimensional soil salinity variation using machine learning algorithms in arid and semi-arid lands of Qazvin Plain. *Iranian Journal of Soil and Water Research*, *52*(7), 1915-1929. https://doi.org/10.22059/ijswr.2021.323030.668957
33. Mousavi, S.R., Sarmadian, F., Angelini, M.E., Bogaert, P., & Omid, M. (2023). Cause-effect relationships using structural equation modeling for soil properties in arid and semi-arid regions. *Catena*, *232*, 107392. https://doi.org/ 10.1016/j.catena.2023.107392
34. Mousavi, S.R., Parsayi, F., Rahmani, A., Sedri, M.H., & Kohsar Bostani, M. (2020). Spatial prediction some of the surface soil properties using interpolation and machine learning models. *Journal of Soil Management and Sustainable Production*, *10*(3), 27-49. (In Persian with English abstract). https://doi.org/10.22069/EJSMS .2021.17251.1916
35. Mousavi, S.R., Sarmadian, F., Rahmani, A., & Khamoshi, S.E. (2019). Digital soil mapping with regression tree classification approaches by RS and geomorphometry covariate in the Qazvin Plain, Iran. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 773-777.
36. Ließ, M., Glaser, B., & Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture: comparison of regression tree and Random Forest models. *Geoderma*, *170*, 70-79. https://doi.org/10.1016/j.geoderma.2011.10.010
37. Organization of Geology and Mineral Explorations of Ira, (1995). Geology map (1:100000) scale. Marvdasht, Fars, Iran.
38. Ostovari, Y., Moosavi, A.A., Mozaffari, H., & Pourghasemi, H.R. (2021). RUSLE model coupled with RS-GIS for soil erosion evaluation compared with T value in Southwest Iran. *Arabian Journal of Geosciences*, *14*, 1-15. https:// doi.org/10.1007/s12517-020-06405-4
39. Olaya, V. I. C. T. O. R. . A gentle introduction to SAGA GIS. The SAGA User Group eV, Gottingen, Germany, 208.
40. Padarian, J., Minasny, B., & McBratney, A.B. (2019). Machine learning and soil sciences: A review aided by machine learning tools. *SOIL*, *6*, 35-52. https://doi.org/10.5194/soil-6-35-2020.
41. Pahlavan-Rad, M.R., & Akbarimoghaddam, A. (2018). Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *Catena*, *160*, 275-281. https://doi.org/10.1016/j.catena.2017.10.002
42. Paramasivam, C.R. (2019). Merits and demerits of GIS and geostatistical techniques. *GIS and Geostatistical Techniques for Groundwater Science*, 17-21.
43. Poppiel, R.R., Lacerda, M.P., Demattê, J.A., Oliveira Jr, M.P., Gallo, B.C., & Safanelli, J.L. (2019). Pedology and soil class mapping from proximal and remote sensed data. *Geoderma*, *348*, 189-206. https://doi.org/10.1016/ j.geoderma.2019.04.028
44. Parent, E.J., Parent, S.É., & Parent, L.E. (2021). Determining soil particle-size distribution from infrared spectra using machine learning predictions: Methodology and modeling. *Plos One*, *16*(7), e0233242. https://doi.org/10.1371/journal.pone.0233242
45. Radočaj, D., Jurišić, M., Antonić, O., Šiljeg, A., Cukrov, N., Rapčan, I., Plaščak, I., & Gašparović, M. (2022). A multiscale cost–benefit analysis of digital soil mapping methods for sustainable land management. *Sustainability*, *14*(19), 12170. https://doi.org/10.3390/su141912170
46. Riza, S., Sekine, M., Kanno, A., Yamamoto, K., Imai, T., & Higuchi, T. (2021). Modeling soil landscapes and soil textures using hyperscale terrain attributes. *Geoderma*, *402*, 115177.
47. Rossel, R.V., & McBratney, A.B. (2008). Diffuse reflectance spectroscopy as a tool for digital soil mapping. In *Digital soil mapping with limited data* (pp. 165-172). Dordrecht: Springer Netherlands. https://doi.org/ 10.1007/978-1-4020-8592-5_13.
48. Sahraei, N., Landi, A., & Hojati, S. (2022). Digital mapping of soil texture components in part of Khuzestan plain lands using machine learning models. Iranian Journal of *Soil and Water Research*, 53(10), 2261-2276. https://doi.org/ 10.22059/ijswr.2022.348442.669360

49. Shahriari, M., Delbari, M., Afrasiab, P., & Pahlavan-Rad, M.R. (2019). Predicting regional spatial distribution of soil texture in floodplains using remote sensing data: A case of southeastern Iran. *Catena*, *182*, 104149. https://doi.org/10.1016/j.catena.2019.104149
50. Sahraei, N., Landi, A., & Hojati, S. (2022). Digital mapping of soil texture components in part of Khuzestan plain lands using machine learning models. *Iranian Journal of Soil and Water Research*, *53*(10), 2261-2276. (In Persian with English abstract). https://doi.org/10.22059/ijswr.2022.348442.669360
51. Sørensen, H. (2004). RPD revisited – a mean to distinguish between poor and good predictions. *Journal of Near Infrared Spectroscopy*, *12*(6), 321-327.
52. Swain, S.R., Chakraborty, P., Panigrahi, N., Vasava, H.B., Reddy, N.N., Roy, S., Majeed, I., & Das, B.S. (2021). Estimation of soil texture using Sentinel-2 multispectral imaging data: An ensemble modeling approach. *Soil and Tillage Research*, *213*, 105134. https://doi.org/10.1016/j.still.2021.105134
53. Taghizadeh-Mehrjardi, R., Toomanian, N., Khavaninzadeh, A. R., Jafari, A., & Triantafilis, J. (2016). Predicting and mapping of soil particle-size fractions with adaptive neuro-fuzzy inference and ant colony optimization in central I ran. *European Journal of Soil Science*, *67*(6), 707-725. https://doi.org/10.1111/ejss.12382
54. Tashayo, B., Honarbakhsh, A., Akbari, M. & Eftekhari, M. (2020). Land suitability assessment for maize farming using a GIS-AHP method for a semi-arid region, Iran. *Journal of the Saudi Society of Agricultural Sciences*, *19*(5), 332-338. https://doi.org/10.1016/j.jssas.2020.03.003
55. Tümsavaş, Z., Tekin, Y., Ulusoy, Y., & Mouazen, A.M. (2019). Prediction and mapping of soil clay and sand contents using visible and near-infrared spectroscopy. *Biosystems Engineering*, *177*, 90-100. https://doi.org/10.1016/j.biosystemseng.2018.06.008
56. Wadoux, A.M.C., Minasny, B., & McBratney, A.B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, *210*, 103359. https://doi.org/10.1016/j.earscirev.2020.103359
57. Wallach, D., Makowski, D., Jones, J.W., & Brun, F. (2006). *Working with dynamic crop models: evaluation, analysis, parameterization, and applications*. Elsevier.
58. Wang, Z., Shi, W., Zhou, W., Li, X., & Yue, T. (2020). Comparison of additive and isometric log-ratio transformations combined with machine learning and regression kriging models for mapping soil particle size fractions. *Geoderma*, *365*, 114214. https://doi.org/10.1016/j.geoderma.2020.114214
59. Wilding, L.P. (1985). *Spatial variability: its documentation, accommodation and implication to soil surveys*. In: Soil Spatial Variability, Las Vegas NV, pp. 166–194.
60. Zeraatpisheh, M., Ayoubi, S., Jafari, A., Tajik, S., & Finke, P. (2019). Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma*, *338*, 445-452. https://doi.org/10.1016/j.geoderma.2018.09.006
61. Ziadat, F.M., 2005. Analyzing digital terrain attributes to predict soil attributes for a relatively large area. *Soil Science Society of America Journal*, *69*, 1590–1599. https://doi.org/10.2136/sssaj2003.0264
62. Zinck, J.A., Metternicht, G., Bocco, G., & Del Valle, H.F. (2015). *Geopedology: An integration of geomorphology and pedology for soil and landscape studies*. Springer.
63. Zhang, Y.Y., Wu, W., & Liu, H. (2019). Factors affecting variations of soil pH in different horizons in hilly regions. *Plos One*, *14*(6), e0218563. https://doi.org/10.1371/journal.pone.0218563
64. Zhang, X., Zhang, W.C., Wu, W., & Liu, H.B. (2023). Horizontal and vertical variation of soil clay content and its controlling factors in China. *Science of The Total Environment*, *864*, 161141.